BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

ing the coding region for the polypeptide. Where chromosomal DNA is employed, the region containing the coding region may be detected employing probes, restriction mapped, and by appropriate techniques isolated substantially free of untranslated 5' and 3' regions. Where only portions of the coding sequence are obtained, the remaining portions may be provided by synthesis of adapters which can be ligated to the coding portions and provide for convenient termini for ligation to other sequences providing particular functions or properties.

Where the two genes are obtained in-whole or in-part from naturally occurring sources, it will be necessary to ligate the two genes in proper reading frame. If cleavage of the fused protein is required, where their juncture does not define a selectable cleavage site, genes will be separated by a selectively cleavable site. The selectively cleavable site will depend to some degree on the nature of the genes. That is, the means for cleaving may vary depending upon the amino acid sequence of one or both genes.

Alternatively, there will be situations where cleavage is not necessary and in some situations undesirable. Fused proteins may find use as diagnostic reagents, in affinity columns, as a source for the determination of a sequence, for the production of antibodies using the fused protein as an immunogen, or the like.

The two genes will normally not include introns, since splicing of mRNA is not extensively employed in the eukaryotic unicellular microorganisms of interest.

The polypeptide of interest may be any polypeptide, either naturally occurring or synthetic, derived from prokaryotic or eukaryotic sources. Usually, the polypeptide will have at least 15 amino acids (gene of 45 bp), more usually 30 amino acids (gene of 90 bp), and may be 300 amino acids (gene of 900 bp) or greater.

Polypeptides of interest include enzymes, viral proteins (e.g. proteins from AIDS related virus, such as p18, p25, p31, gp41, etc.), mammalian proteins, such as those involved in regulatory functions, such as lymphokines, growth factors, hormones or hormone precursors (e.g., proinsulin, insulin like growth factors, e.g., IGF-I and -II, etc.), etc., blood clotting factors, clot degrading factors, immunoglobulins, etc. Fragments or fractions of the polypeptides may be employed where such fragments have physiological activity, e.g., immunological activity such as cross-reactivity with the parent protein, physiological activity as an agonist or antagonist, or the like.

One of the methods for selectable cleavage is cyanogen bromide which is described in U.S. Pat. No. 4,366,246. This technique requires the absence of an available methionine other than at the site of cleavage or the ability to selectively distinguish between the methionine to be cleaved and a methionine within the polypeptide sequence. Alternatively, a protease may be employed which recognizes and cleaves at a site identified by a particular type of amino acid. Common proteases include trypsin, chymotrypsin, pepsin, bromelain, papain, or the like. Trypsin is specific for basic amino acids and cleaves on the carboxylic side of the peptide bond for either lysine or arginine. Further, peptidases can be employed which are specific for particular sequences of amino acids, such as those peptidases which are involved in the selective cleavage of secretory leader signals from a polypeptide. These enzymes are specific for such sequences which are found with α-factor and killer toxin in yeast, such as *KEX* 2 endopeptidase with specificity for pairs of basic residues (Julius et al., *Cell* (1984) 37:1075–1089). Also, enzymes exist which cleave at specific sequences of amino acids. Bovine enterokinase (Light et al., *Anal. Biochem.* (1980) 106:199–206) cleaves to the carboxylic side of lysine or arginine that is preceded by acid residues of aspartic acid, glutamic acid, or carboxymethyl cysteine. Particularly useful is the sequence $(Asp)_4$ Lys found naturally as part of the activation peptide of trypsinogen in many species. Other enzymes which recognize and cleave specific sequences include: Collagenase (Germino and Batia, *Proc. Natl. Acad. Sci.* (1984) 81:4692–4696); factor X (Nagai & Thygersen, *Nature* (1984) 309:810–812); and polyubiquitin processing enzyme (Ozakaynak et al., *Nature* (1984) 312:663–666).

In addition to the amino acids comprising the cleavable site, it may be advantageous to separate further the two fused polypeptides. Such a "hinge" would allow for steric flexibility so that the fused polypeptides would be less likely to interfere with each other, thus preventing incorrect folding, blockage of the cleavage site, or the like.

The "hinge" amino acid sequence could be of variable length and may contain any amino acid side chains so long as the side chains do not interfere with the mode of action employed to break at the cleavable site or with required interactions in either fused polypeptide, such as ionic, hydrophobic, or hydrogen bonding. Preferably the amino acids comprising the hinge would have side chains that are neutral and either polar or nonpolar and may include one or more prolines. The hinge region will have at least one amino acid and may have 20 or more amino acids, usually not more than 15 amino acids, particularly the nonpolar amino acids G, A, P, V, I, L, and the neutral polar amino acids, N, Q, S, and T.

Exemplary hinge sequences may be, but are not limited to: N-S; Q-A; N-S-G-S-P; A-A-S-T-P; N-S-G-P-T-P-P-S-P-G-S-P; S-S-P-G-A; and the like. It is contemplated that such hinge sequences may be employed as repeat units to increase further the separation between the fused polypeptides.

So that the "hinge" amino acids are not bound to the final cleaved polypeptide of interest, it is desirable, but not required to practice the invention, to place the "hinge" between the polypeptide that is produced independently at high yield and the sequence for the cleavable site.

Where one or more amino acids are involved in the cleavage site, the codons coding for such sequence may be prepared synthetically and ligated to the sequences coding for the polypeptides so as to provide for a fused protein where all the codons are in the proper reading frame and the selectable cleavage site joins the two polypeptides.

Instead of only a small portion of the fused coding sequence being synthetically prepared, the entire sequence may be synthetically prepared. This allows for certain flexibilities in the choice of codons, whereby one can provide for preferred codons, restriction sites, avoid or provide for particular internal structures of the DNA and messenger RNA, and the like.

While for the most part, the fused coding sequence will be prepared as a single entity, it should be appreciated that it may be prepared as various fragments, these fragments joined to various untranslated regions, providing for particular functions and ultimately the coding sequences brought together at a subsequent stage. However, for clarity of presentation, the discussion will

zyme XhoI, treated with nuclease S1, and subsequently treated with the restriction enzyme EcoRI to create a linear vector molecule having one blunt end at the site of the XhoI linker and an EcoRI end.

The GAP portion of the promoter was constructed by cutting plasmid pPGAP (supra) with the enzymes BamHI and EcoRI, followed by the isolation of the 0.4Kbp DNA fragment. The purified fragment was cut with the enzyme AluI to create a blunt end near the BamHI site.

Plasmid pJS104 was constructed by the ligation of the AluI-EcoRI GAP promoter fragment to the ADH2 fragment present on the linear vector described above.

Plasmid pJS104 was digested with BamHI (which cuts upstream of the ADH2 region) and with NcoI (which cuts downstream of the GAP region). The about 1.3Kbp fragment containing the ADH2-GAP promoter was gel purified and ligated to an about 1.7Kbp fragment containing the hSOD-proinsulin fusion DNA sequences and GAP terminator present in pYSI1 (previously described). This 3Kbp expression cassette was cloned into BamHI digested and phosphatase treated pC1/1 to yield pYASI1.

Construction of pYASI1 Derivatives Containing Trypsin and Enterokinase Cleavage Sites

A series of plasmids were constructed derived from pYASI1, in which the GAP terminator was replaced by the α-factor terminator (Brake et al., *Proc. Natl. Acad. Sci. USA* (1984) 81:4642) and the cleavage site between SOD and proinsulin was modified to code for trypsin or enterokinase processing sites. Sequences coding for Lys-Arg were used to replace the methionine codon in pYASI1 yielding a trypsin site. Alternatively, sequences coding for $(Asp)_4Lys$ were used at the cleavage site to yield an enterokinase site. In addition, sequences coding for extra hinge amino acids were also inserted between the SOD and the cleavage site in other constructions.

Expression of Fusion Proteins

Yeast strain 2150-2-3 (*Mat a, ade* 1, leu 2-04, cir°) or P017 (Mat a, leu 2-04, cir°) were transformed with the different vectors according to Hinnen et al., *Proc. Natl. Acad. Sci. USA* (1978) 75:1929-1933. Single transformant colonies harboring constitutive GAP regulated vectors were grown in 2 ml of leu⁻ selective media to late log or stationary phase. Cells harboring inducible ADH2-GAP regulated vectors were grown to saturation in leu⁻ selective media, subsequently diluted 1:20 (v/v) in YEP, 3% ethanol, with or without 2-3.5 mM $CuSO_4$ and grown to saturation in this medium. Cells were lysed in the presence of SDS and reducing agent and the lysates clarified by centrifugation. Cleared lysates were subjected to polyacrylamide gel electrophoresis (Laemmli, *Nature* (1970) 277:680). Following staining with Coomassie blue, a band of about 28 kDal (kilodaltons) was observed, the size predicted for the fusion protein. This band was detected in those cells transformed with expression vectors, while being absent from extracts of cells harboring control (pC1/1) plasmids. Amount of protein per band was determined by densitometric analysis of the Coomassie blue stained gels. The fusion protein accounts for over 10% of the total cell protein as estimated from the stained gels in those cells transformed with pYSI1, pYSI2 or pYASI1, while it accounts for less than 0.5% in pYPKI1 or pYPKI2 transformants (See Table 1).

TABLE 1

The Yield of SOD-PI from 2150 or P017 Transformed with Different Expression Plasmids and Grown in the Absence/Presence of 2-3.5 mM $CuSO_4$

Strain	Plasmid	Description of sequences contained in the expression cassette.	Expression (percent of total cell protein)[1] −Cu++	+Cu++
2150	pYPKI1	PYK_p PYK M BCA5 GAP_t	0.5	
2150	pYPKI2	GAP_p M BCA5 KRSTS $_2$PYK PYK_t	0.5	
2150	pYSI1	GAP_p SOD M BCA5 GAP_t	10	
2150	pYSI2	GAP_p M BCA5 $KR(ST)_2S$ SOD GAP_t	10	
		SOD Met Proinsulin		
2150	pYASI1	$(ADH—GAP)_p$ SOD M BCA5 GAP_t	10	
P017	pYASI1	$(ADH—GAP)_p$ SOD M BCA5 GAP_t	20-30	20-30
		SOD (hinge) $(Asp)_4$LysPI		
P017	pYSI12	$(ADH—GAP)_p$SOD—D_4K—BCA5 α-factor$_t$	6-9	11-14
P017	pYSI15	$(ADH—GAP)_p$SOD—(NS)D_4K—BCA5 α-factor$_t$	5-6	9-14
P017	pYSI8	$(ADH—GAP)_p$SOD—(NSGSP)D_4K—BCA5 α-factor$_t$	5	8
P017	pYSI4	$(ADH—GAP)_p$SOD—(NSGPTPPSPGSP)D_4K—BCA5 α-factor$_t$	9-12	9-16
		SOD (hinge) LysArgPI		
P017	pYSI13	$(ADH—GAP)_p$SOD—KR—BCA5 α-factor$_t$	8-10	8-10
P017	pYSI10	$(ADH—GAP)_p$SOD—(NSGSP)KR—BCA5 α-factor$_t$	5-7	10-15
P017	pYSI3	$(ADH—GAP)_p$SOD—(NSGPTPPSPGSP)KR—BCA5 α-factor$_t$	5-8	15-30

[1]Determined by scanning densitometer analysis of Coomassie Blue stained gels.

Note:

Proinsulin (PI) accounts for less than 0.1% of total cell protein in cells transformed with pYGAPINS5, a plasmid containing the proinsulin gene under regulation of GAPDH promoter and terminator (GAP_p M BCA5 GAP_t).

PYK: pyruvate kinase gene
SOD: human SOD gene
BCA5: proinsulin gene
P, G, D, N, M, K, R, S, T: one letter amino acid code
GAP_p: GAP promoter
GAP_t: GAP terminator
PYK_p: PYK promoter
PYK_t: PYK terminator
$(ADH2—GAP)_p$: hybrid
ADH2—GAP: promoter
α-factor$_t$: α-factor terminator

Results shown in Table 1 indicate that while expression levels of PYK-proinsulin fusion are comparable to those obtained with proinsulin alone (about 0.5% and 0.1%, respectively), the expression levels of hSOD-proinsulin are about 20 to 100 fold higher. The inducible ADH2-GAP hybrid transcriptional initiation regula-

twice against 10 mM Tris, pH 7.5, and once against 1 mM Tris.

The product, ~90% pure proinsulin-S-sulfonate, was shown to migrate as expected on pH 9 gel electrophoresis (Linde et al., *Anal. Biochem.* (1980) 107:165–176), and has the correct 15 N-terminal residues. On analysis, the amino acid composition was very close to that expected, not exactly correct due to the presence of a low level of impurities. The yield was 150 mg.

Preliminary results on renaturation have been obtained with the following procedure. The proinsulin-S-sulfonate can be renatured at pH 10.5, with β-mercaptoethanol (Frank et al. (1981) in Peptides: Synthesis, Structure and Function, Proceedings of the Seventh American Peptide Symposium, Rich and Gross, eds., Pierce Chemical Co., Rockford, Ill., pp. 729–738). In preliminary experiments, the yield of correctly renatured proinsulin has been monitored by the production of insulin produced from digestion with trypsin and carboxypeptidase B. The proinsulin —S—SO_3 produced by this process appears to renature as well as purified porcine proinsulin —S—SO_3. This process has been reported to yield 70% of the expected amount of insulin. The insulin produced in this way has the correct N-terminal 15 residues of each A chain and B chain as determined by amino acid sequencing.

EXAMPLE II

Construction and Expression of Expression Vectors for SOD-p31 Fusion Protein

A yeast expression plasmid pC1/1-pSP31-GAP-ADH2, containing the human SOD gene fused to the amino terminus or the endonuclease region (p31) of the pol gene of the AIDS related virus (ARV) (Sanchez-Pescador et al., *Science* (1985) 227:484) was constructed. Expression of SOD-p31 is non-constitutive and is under regulation of a hybrid ADH-GAP promoter.

Construction of pC1/1-pSP31-GAP-ADH2 Derivative

For the construction of a gene for a fused protein SOD-p31 to be expressed in yeast, a plasmid (pS14/39–2) was used. This plasmid contains the SOD gene fused to the proinsulin gene under the regulation of the ADH-2/GAP promoter in the same manner as pYAS1. The proinsulin gene is located between EcoRI and SalI restriction sites. To substitute the proinsulin gene with the p31 fragment, two oligomers designated ARV-300 and ARV-301, respectively, were synthesized using phosphoramidite chemistry. The sequences generate cohesive ends for EcoRI and NcoI on each side of the molecule when the two oligomers are annealed. ARV-300 and ARV-301 have the sequences:

ARV-300 5' AATTCAGGTGTTGGAGC
GTCCACAACCTCGGTAC 5' ARV-301

Two μg of pS14/39–2 linearized with EcoRI were ligated to 100 picomoles each of phosphorylated ARV-300 and dephosphorylated ARV-301 in the presence of ATP and T4 DNA ligase in a final volume of 35 μl. The reaction was carried out at 14° C. for 18 hr. The DNA was further digested with SalI and the fragments were resolved on a 1% low melting point agarose gel and a fragment containing the vector plus the SOD gene (~6.5 kb) was purified as described above and resuspended in 50 μl of TE (10 mM Tris, 1 mM EDTA, pH 8). Five μl of this preparation were ligated to 5 μl of the p31 fragment (ARV248NL, see below) in 20 μl final

volume for 18 hr at 14° C. and 5 μl used to transform competent HB101 cells. The resultant plasmid was called pSP31. Twenty μg of this plasmid were digested with BamHI and a fragment of about 2900 bp was isolated by gel electrophoresis, resuspended in TE and ligated to pC1/1 previously cut with BamHI. This DNA was used to transform HB101 and transformants with the BamHI cassette were obtained. Yeast strain P017 (Mat a, leu2-04, cir°) was transformed with this pC1/1-pSP31-GAP-ADH2 derivative.

Preparation of ARV248NL, the p31 Coding Fragment.

The 800 bp ARV248NL fragment codes for numbered amino acids 737 to the end of the pol protein as shown in FIG. 2 of Sanchez-Pescador et al. (supra). The following procedure was used for its preparation.

A 5.2 kb DNA fragment was isolated from a KpnI digest of ARV-2 (9B) (Sanchez-Pescador et al., supra) containing the 3' end of the pol gene, orf-1, env and the 5' end of orf-2, that had been run on a 1% low melting point agarose (Sea-Pack) gel and extracted with phenol at 65° C., precipitated with 100% ethanol and resuspended in TE. Eight μl of this material were further digested with SstI for 1 hr at 37° C. in a final volume of 10 μl. After heat inactivation of the enzyme, 1.25 μl of this digest were ligated to 20 ng of M13mp19 previously cut with KpnI and SstI, in the presence of ATP and in a final volume of 20 μL The reaction was allowed to proceed for 2 hr at room temperature. Five μl of this mixture were used to transform competent *E. coli* JM101. Clear plaques were grown and single-stranded DNA was prepared as described in Messing and Vieira, *Gene* (1982) 19:269–276.

The DNA sequence in the M13 template was altered by site specific mutagenesis to generate a restriction site recognized by NcoI (CCATGG). An oligodeoxynucleotide that substitutes the A for a C at position 3845 (FIG. 1 in Sanchez-Pescador et al., supra) and changes a T for an A at position 3851 was synthesized using solid phase phosphoramidite chemistry. Both of these changes are silent in terms of the amino acid sequence, and the second one was introduced to decrease the stability of the heterologous molecules. The oligomer was named ARV-216 and has the sequence:

5'-TTAAAATCACTTGCCATGGCTCT-
CCAATTACTG

and corresponds to the noncoding strand since the M13 derivative template 01100484 is single-stranded and contains the coding strand. The 5' dephosphorylated M13 sequencing primer, 50 mM Tris-HCl pH 8, 20 mM KCl, 7 mM $MgCl_2$ and 0.1 mM EDTA. The polymerization reaction was done in 100 μl containing 50 ng/μl DNA duplex, 150 μM dNTPs, 1 mM ATP, 33 mM Tris-acetate pH 7.8, 66 mM potassium acetate, 10 mM magnesium acetate, 5 mM dithiothreitol (DTT), 12.5 units of T4 polymerase, 100 μg/ml T4 gene 32 protein and 5 units of T4 DNA ligase. The reaction was incubated at 30° C. for 30 min and was stopped by the addition of EDTA and SDS (10 mM and 0.2% respectively, final concentration). Competent JM101 *E. coli* cells were transformed with 1, 2, and 4 μl of a 1:10 dilution of the polymerization product and plated into YT plates. Plaques were lifted by adsorption to nitrocellulose filters and denatured in 0.2N NaOH, 1.5M NaCl, followed by neutralization in 0.5M Tris-HCl pH 7.3, 3M NaCl and equilibrated in 6×SSC. The filters were blotted dry, baked at 80° C. for 2 hr and preannealed at 37° C. in 0.2% SDS, 10× Denhardt's 6×SSC. After 1 hr, 7.5×106 cpm of labelled ARV-216 were added to the filters and incubated for 2 additional hr at 37° C. The filters were washed in 6×SSC at 42° C. for 20 min, blot-dried and used to expose film at −70° C. for 1 hr using an intensifying screen. Strong hybridizing plaques were grown and single-stranded DNA was prepared from them and used as templates for sequencing. Sequencing showed that template 01021785 contains the NcoI site as well as the second substitution mentioned above.

A second oligomer was synthesized to insert sites for SalI and EcoRI immediately after the termination codon of the pol gene (position 4647, FIG. 1, Sanchez-Pescador et al., supra). This oligomer was called ARV-248 and has the sequence:

5'-GGTGTTTTACTAAAGAATTCCGTCGAC-
TAATCCTCATCC.

Using the template 01020785, site specific mutagenesis was carried out as described above except that the filter wash after the hybridization was done at 65° C. As above, 8 strong hybridizing plaques were grown and single-stranded DNA was sequenced. The sequence of template 01031985 shows that it contains the restriction sites for NcoI, SalI, and EcoRI as intended.

Replicative form (RF) of the M13 01031098 template was prepared by growing 6 clear plaques, each in 1.5 ml of 2×YT (0.5% yeast extract, 0.8% tryptone, 0.5% HaCl, 1.5% agar) at 37° C. for 5 hr. Double-stranded DNA was obtained as described by Maniatis, et al., *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor, 1982, pooled and resuspended in 100 μl final volume. A 20 μl aliquot of RF was cut with NcoI and SalI in a 40 μl volume of digestion buffer. This fragment was used for p31 expression in yeast. The samples were run on a 1% low melting point agarose (Sea-Pack) gel and the DNAs were visualized by fluorescence with ethidium bromide. The 800 bp band was cut and the DNA was extracted from the gel as mentioned above and resuspended in 10 μl of TE. The fragment was called ARV248NL.

Induction of pC1/1-pSP31-GAP-ADH2

Three different kinds of inductions were tried:

(1) P017 colonies were induced in either a 10 ml culture of YEP/1% glucose or a leu−/3% ethanol culture for 24 hr. The yeast pellets from each mixture were analyzed for p31 by both polyacrylamide gels and Westerns using sera from AIDS patients. Even though the Coomassie-stained gel showed a negative result, in both cases the Western did light up a band of the correct molecular weight.

(2) P017 colonies were induced in a 30 ml culture of YEP/1% ethanol for 48 hr. Aliquots were analyzed by PAGE at various time points during the induction. The Coomassie-stained gel shows a band in the correct molecular weight range (47–50 kd) that appears after 14 hr in YEP/1% ethanol and reaches a maximum intensity at 24 hr of induction. The Western result for SOD p31 using sera from AIDS patients correlates well with the Coomassie-stained gel, showing strong bands at 24 and 48 hr of induction.

Purification and Characterization of SOD-p31 from Yeast

Frozen yeast (bacteria) cells were thawed at room temperature and suspended in 1.5 volumes of lysis buffer (20 mM Tris-Cl, pH 8.0, 2 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), for bacteria; 50 mM Tris-Cl, pH 8.0, 2 mM EDTA, 1 mM PMSF for

yeast), and mixed with 1 volume of acid-washed glass beads.

Cells were broken for 15 min. in a non-continuous mode using the glass chamber of a Dynomill unit at 3,000 rpm, connected to a −20° C. cooling unit. Glass beads were decanted for 2–3 min. on ice, and the cell lysate removed. The decanted glass beads were washed twice with 30 ml of lysis buffer at 4° C. The cell lysate was centrifuged at 39,000×g for 30 min.

The pellet obtained from the above centrifugation was washed once with lysis buffer, after vortexing and suspending it at 4° C. (same centrifugation as above). The washed pellet was treated with 0.2% SDS (for bacteria) and 0.1% SDS (for yeast) in lysis buffer and was agitated by rocking at 4° C. for 10 min. The lysate was centrifuged at 39,000×g for 30 min. The pellet was boiled in sample buffer (67.5 mM Tris-Cl, pH 7.0, 5% β-mercaptoethanol, 2.3% SDS) for 10 min and centrifuged for 10 min. at 39,000×g. The supernatant was recovered and passed through a 0.45 μm filter. The supernatant from the above filter was loaded (maximum 50 mg of protein) on a gel filtration column (2.5×90 cm, ACA 34 LKB) with a flow rate of 0.3–0.4 ml/min, equilibrated with phosphate-buffered saline (PBS), 0.1% SDS. The fractions containing SOD-p31 were pooled and concentrated either by vacuum dialysis or using a YM5 Amicon membrane at 40 psi. The protein was stored at −20° C. as concentrated solution.

Gel electrophoresis analysis showed that the SOD-p31 protein migrates having a molecular weight of about 46 kd and is over 90% pure.

Similar constructions and results have been obtained by expressing an SOD-p31 fusion under regulation of a bacterial trp-lac promoter in *E. coli*.

The SOD-p31 fused protein finds use in immunoassays to detect the presence of antibodies against AIDS in body fluids. Successful results have been obtained using the SOD-p31 fusion protein in ELISA as well as in strip assays.

EXAMPLE III

Construction and Expression of Expression Vectors for SOD-IGF-2 Fusion Protein

A yeast expression plasmid pYLUIGF2–14, containing the human SOD gene fused to the amino terminus of the IGF2 gene (see EPO 123 228) was constructed. Expression of SOD-IGF2 is non-constitutive and it is under regulation of a hybrid ADH-GAP promoter.

Construction of pYLUIGF2–14

For the construction of a gene for a fused protein SOD-IGF2 to be expressed in yeast, plasmid pYS18 was used. Plasmid pYS18 contains the SOD gene fused to the proinsulin gene under the regulation of the ADH-GAP promoter and α-factor terminator (see Table 1). Plasmid pYS18 was digested with BamHI and EcoRI. The 1830 bp fragment (containing the ADH-GAP promoter and SOD gene) was purified by gel electrophoresis.

A second BamHI (460 bp) fragment coding for amino acid residues 41 to 201 of IGF-2 and for the α-factor terminator (see EPO 123 228) was ligated to the following linker:

EcoRI SalI

AATTCCATGGCTTACAGACCATCCGAAACCTTGTGTGGTGGTGAATTGG

GGTACCGAATGTCTGGTAGGCTTTGGAACACACCACCACTTAACCAGCT

The linker provides for an EcoRI overhang, an ATG codon for methionine and for codons 1–40 of IGF2 and SaLI overhang.

The resulting EcoRI-BamHI (510 bp fragment containing the IGF-2 gene and α-factor terminator was ligated to the 1830 bp BamHI-EcoRI fragment containing the ADH-GAP promoter and SOD (see above). The resulting BamHI (2340 bp) fragment was cloned into BamHI digested and phosphatase treated pAB24 (see below) to yield pYLUIGF2–14.

pAB24 is a yeast expression vector (see FIG. 2) which contains the complete 2μ sequences (Broach 1981, In: Molecular Biology of the Yeast Saccharomyces 1:445, Cold Spring Harbor Press) and pBR322 sequences. It also contains the yeast URA3 gene derived from plasmid YEp24 (Botstein et al., 1979 *Gene* 8:17) and the yeast $LEU2^d$ gene derived from plasmid pC1/1 (see EPO 116201). Insertion of the expression cassette was in the BamHI site of pBR322, thus interrupting the gene for bacterial resistance to tetracycline.

Expression of SOD-IGF2

Yeast AB110 (Matα, ura3-52, leu2-04 or both leu2-3 and leu2-112, pep4-3, his4-580, cir°) was transformed with pYLUIGF2–14. Transformants were grown up on ura⁻ selective plate. Transformant colonies were transferred to 3 ml leu⁻ selective media and grown 24 hrs in 30° C. shaker. 100 μl of a 1×10^{-4} this culture was plated onto ura⁻ plates and individual transformants were grown up for ~48–72 hrs. Individual transformants were transferred to 3 ml leu⁻ media and grown 24 hrs in a 30° C. shaker. One ml each of these cultures was diluted into 24 ml UEP, 1% glucose media and cells were grown for 16–24 hours for maximum yield of SOD-IGF2. Cells were centrifuged and washed with H_2O. Cells were resuspended in 2-volumes of lysis buffer (phosphate buffer, pH 7.3 (50–100 mM), 0.1% Triton X100). Two volumes of acid washed glass beads were added and the suspension was alternately vortexed or set on ice (5×1 min. each cycle). The suspension was centrifuged and the supernatant decanted. The insoluble pellet was incubated in lysis buffer 1% SDS at room temperature for 30 min. The suspension was centrifuged and the supernatant was frozen and lyophilized.

Two other constructions: pYLUIGF2–15 and pYUIGF2–13 were used as controls for expression of a non-fused IGF2. The former plasmid (pYLUIGF2–15) for intracellular expression contains the IGF2 gene under control of the GAP promoter and α-factor terminator. The latter plasmid (pYUIGF2-13) for secretion of IGF2, the IGF-2 gene under control for the GAP promoter, α-factor leader and α-factor terminator.

	EXPRESSION OF IGF2	
Construction in AB110	PAGE STAIN (% of total cell protein)	RRA+
1. pYLUIGF2-15 ($GAP_pIGF2\text{-}\alpha f_t$)	NOT DETECTABLE	NA
2. pYUIGF2-13 ($GAP_p\alpha f_L\text{-}IGF2\text{-}\alpha f_T$)	BARELY DETECTABLE	10 μg/l
3. pYLUIGF2-14	10–15%	NA*

United States Patent [19]
Cousens et al.

[11] Patent Number: 4,751,180
[45] Date of Patent: Jun. 14, 1988

-continued

Construction in AB110	EXPRESSION OF IGF2 PAGE STAIN (% of total cell protein)	RRA+
(ADH2/GAP$_\rho$SOD-IGF2-α[7)		

NA: Not available.
*By Coomassie blue staining, the SOD-IGF2 fusion protein represents 10–15% of the total cell protein, i.e., ~100–300 mg/l culture equivalent. IGF2 represents ~½ of the fusion protein, therefore it constitutes about 30–100 mg/l culture equivalent. Analytical CNBr cleavage reactions with the fusion protein have resulted in a band on PAGE which migrates to the position expected for IGF2.
+RRA-IGF2 levels were measured by a placental membrane radioreceptor assay (RRA) according to Homer et al., J. of Clinical Endocrinology and Metabolism (1978) 47:1287 and Marshall et al., J. of Clinical Endocrinology and Metabolism (1974) 39:283. Placental membranes for the RRA were prepared by the method of Cuatrecasas, Proc. Natl. Acad. Sci. USA (1972) 69:318.

Protocol for CNBr Cleavage of SOD-IGF2

The insoluble fraction from glass bead lysis of yeast cells was dissolved in 70% formic acid. CNBr crystals (~/g CnBr/100 mg fusion protein) were added and incubation was carried out at room temperature for 12–15 h in the dark. This step may be repeated after 24 hrs if cleavage is incomplete.

It is evident from the above results that otherwise difficultly and inefficiently produced polypeptides may be produced in substantially enhanced yields by employing a fused protein, where the fusion protein includes a relatively short stable polypeptide sequence joined to the other polypeptide by a selectively cleavable site. Thus, high levels of the fusion protein are obtained in a eukaryotic host, such as yeast, allowing for the efficient production of desired polypeptides heterologous to the host.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.